INTERPERSONAL ACCURACY, BIAS AND THE SELECTION INTERVIEW

by

Vivienne L. Bull

A thesis submitted in
partial fulfilment of the
requirements for the
degree of Master of
Science in Psychology at
the University of
Canterbury

University of Canterbury
1989

ACKNOWLEDGEMENTS

I would like to thank Dr Garth Fletcher for his support and timely advice throughout the course of this study.

I am extremely grateful to my parents, Beverley and Russell Bull, friends and colleagues, especially Nicky Sinclair and Malcolm Fletcher, for their assistance and support.

I would also like to thank the people who participated in this study as subjects. Without their patience and co-operation, this study would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF APPENDICES

ABSTRACT

The aim of this research was to investigate the interpersonal accuracy of individuals in the context of a selection interview. Interviewers were twenty University Teaching Fellows from departments other than psychology. Interviewees were ten relatively inexperienced tutors from the Psychology Department. Each interviewer conducted two mock interviews with the aim of predicting the job performance of the people she interviewed. Interviewers also completed a self evaluation (measuring the extent to which they were projecting), a stereotype evaluation (to control for stereotype accuracy), as well as the Attributional Complexity Scale (an individual difference measure). All scales (except the Attributional Complexity Scale) were modified forms of questionnaires used as part of tutor feedback in the Psychology Department. Interviewers were rated by their supervisors as well as the students they had tutored during the year. These measures were used as criteria to compare with the predictions of job performance made by the interviewers. As predicted, interviewer subjects who used projection were more accurate in their predictions of job performance than those who didn't. This effect however, was present only in the first interview. As predicted, agreement between the supervisor and the student evaluations about the same interviewee, was significant. However, the results did not support the prediction that accuracy of the interviewers would exceed chance levels. These results are discussed in relation to interpersonal perception and the selection interview.

INTRODUCTION

Prediction of other people's behaviour is an important facet of our lives. An influential body of research conducted in the 1970's raised important questions concerning the perceptive and predictive abilities of the social perceiver.

This body of research examined people's social judgemental abilities from a process perspective; that is, variables affecting reasoning processes were investigated. This style of research is exemplified primarily in the work of Tversky and Kahneman.

A notable feature of these studies concerns the nature of the assumptions held by these investigators. The underlying premise is essentially negative as the null hypothesis used is: *subjects will commit no errors.* Not surprisingly, people do commit errors, although this is not necessarily negative. For example, expert musicians commit errors, yet they are still exceptional artists . Nevertheless, findings from this research led investigators to conclude that human reasoning processes are fundamentally flawed. For example, Nisbett and Borgida (1975) contend that various psychological discoveries have 'bleak implications' for human rationality.

However, more recently the area has witnessed a surge of interest in a new direction. In contrast with earlier bias and error work, investigators are now concentrating more on what people can do, as opposed to their failings (Funder & Colvin, 1988). This partially stems from a dissatisfaction with the methodologies used in past bias/error research, as

well as a discontent with the poor light in which the social perceiver is viewed.

With regard to the first issue, the usual methodology of bias and error studies plucks the perceiver out of any realistic context and places him or her into strictly controlled laboratory conditions. Thus, it is argued that the studies lack external validity. For example, Swann (1984) contends that the context people are in is meaningful and influential; people take cues from their environment, be it their work place or a psychological laboratory.

More recently, some writers (see Funder, 1987; Swann, 1984) have contended that the pervasively negative perspective taken is unrealistic and not necessarily representative of the actual abilities of the social perceiver. Thus, in contrast to the earlier studies focusing on process, there is a new emphasis on the content of social judgements; that is, the outcomes of social perception and their relation to what is actually occurring in the person's social environment.

One example of a realistic context is the selection interview. Interviewers are subject to powerful influences from the needs of the specific organization for which the interviewer is working, as well as the demands of the interview situation itself. For example, an interviewer may have selection quotas and time restrictions to consider, as well as attempting to predict how a certain interviewee is going to perform if he or she is given the job.

Research within the selection interview area has also tended to focus on biases supposedly specific to the situation. Yet, in contrast to research conducted within the bias/error process paradigm in social psychology,

studies tend to be of a practical and functional nature. Further, research in the selection area has tended to ignore work done in the interpersonal perception area, regardless of the fact that interviewing is essentially a social psychological phenomenon.

The present study then, is intended to help redress the balance in both the social interpersonal perception area, and the selection interview field. Thus, I planned to conduct a study based on an interview setting, but to analyze it in terms of the accuracy of the interviewers in predicting potential job performance of people interviewed. To achieve these aims, I arranged pseudo-interviews. As interviewees I used psychology post-graduate students who tutor Stage 1 psychology students in a laboratory class. This group was chosen because all individuals were relatively inexperienced. Moreover, they had received no formal tutoring feedback. Interviewers were Teaching Fellows from departments other than Psychology. These individuals were used because I was assumed that teaching/tutoring/communication skills were universal to any department. Experienced Teaching Fellows would plausibly be able to recognize these abilities in other people. Each interviewer conducted two interviews.

Following the interview, interviewers predicted how well the interviewees would be rated on a measure of job performance should they be hired. Measures of actual job performance were obtained from interviewees' supervisors as well as from the students they tutored. Interviewers were also required to complete a self evaluation (to measure the extent to which they were using estimations of their own tutoring abilities in their predictions of interviewees' tutoring abilities), a

stereotype evaluation of how they thought the average tutor would perform (to control for stereotype accuracy), and the Attributional Complexity Scale (an individual difference measure).

The following introductory section will be presented in nine major parts. First, I will briefly review the history of social psychological research in interpersonal perception, commenting on the waning of the bias/error process paradigm. I shall also comment on the resurgence of interest in the measurement of accuracy of social judgement. Second, I shall review several recent studies with the aim of investigating whether individuals can achieve a significant level of accuracy in their social judgements. I shall then present research investigating the validities of acquaintances' social judgements. Fourth, I will provide a rationale for the inclusion of an individual differences scale. Fifth, I will look at validity studies conducted within the selection interview area and then move on to study the conclusions from bias research carried out in the selection interview area. Seventh, I shall provide a rationale for my methodology based on Cronbach's conceptual and methodological critique of interpersonal accuracy research (Cronbach, 1955; Gage & Cronbach, 1955). Finally, I shall present the main hypotheses of this study.

History of Interpersonal Perception Research

A pervasive theme in social psychological research has been to assess the competency of the lay perceiver in social judgement. One major controversy centres on the question of whether people are blinded by

ubiquitous biases, or are discerning psychologists who are steadfast adherents to rational logic. One particular branch of research has investigated variables likely to affect the outcome of decisions.

Research on process was never originally intended to directly apply to the issue of content or the outcome of decisions, and researchers concentrated specifically on the former area. Yet, a marked switch in emphasis occurred when investigators began conducting research using normative models. Theories of how people actually think and reason came to be seen more as models of how people should think (Funder, 1987). Suddenly, studies demonstrating how individuals deviate from these idealized ways of reasoning were taken to have disturbing implications for the validity of people's judgements. In short, it appeared that individuals were not very good at social judgement.

This emphasis is exemplified mainly by researchers such as Tversky, Kahneman and Fischoff. Kahneman and Tversky (1973) describe a study demonstrating people's use of the representativeness heuristic. The representativeness heuristic is the tendency to estimate the probability of event B occurring by relying on the degree to which A resembles B.

In this particular study, two groups of subjects were shown brief personality sketches of several people who had been allegedly sampled at random from a group of one hundred individuals, all of whom were either lawyers or engineers. Group 1 were informed that the population from which the personality descriptions were drawn consisted of seventy engineers and thirty lawyers. Group 2 were told that the population consisted of seventy lawyers and thirty engineers. Subjects were required to estimate the probability that a given person was an engineer. Yet,

subjects from both groups gave similar probability judgements, regardless of the fact that in Group 1 the probability of being an engineer was higher and in Group 2 the probability of being a lawyer was higher. The interesting point is that this effect occurred only when personality descriptions of the targets were provided. When no personality sketches were given, it was found that subjects estimated probabilities correctly.

Other prominent examples of these biases are the hindsight phenomenon and the availability heuristic. The hindsight phenomenon is said to occur when judges with knowledge of a particular outcome tend to overestimate the probability they would have declared prior to the event (Fischoff, 1975). The availability heuristic describes the case where the ease with which instances or associations are brought to mind, is an important determinant of frequency or probability (Kahneman & Tversky 1973; Tversky & Kahneman, 1974).

Nevertheless, many aspects of the bias/error process paradigm have been criticized. Further research has tended to raise serious doubts regarding the underlying assumptions of these investigators. Common sense tells us that people are not that bad at making social judgements and can be relatively accurate in complex situations such as predicting who will be good to talk to at a party. The bleak implications this earlier research had for people's social judgemental abilities led Swann (1984) to exclaim:

> The uniformly unflattering image of the social perceiver .... could lead one to wonder how people ever manage to muddle their way through their social relationships. (p. 459)

In a more empirical vein the earlier bias/ error studies have been attacked with regard to using inappropriate criteria in addition to removing the person from a realistic context.

Criterion Problem

As mentioned above, problems arose when normative models of inferential validity were used to show how people should reason; that is they were being used as criteria of social judgement.

Yet it can be shown that changing the definition of bias has dramatic consequences for how competent people appear to be in social judgement. For example, if bias is defined in terms of a 'subjectively based preference' favouring some conclusion over another (Kruglanski & Ajzen, 1983), the implication is that *all* knowledge is biased. Similarly, McArthur and Baron (1983) stress the fundamental accuracy of social perceivers. From an ecological perspective, bias is merely a case of selective attention and action.

However, as described earlier, within the bias/error paradigm bias is a negatively value-laden concept. Yet, according to other definitions it is merely a descriptive concept implying a functional quality to the notion of bias. Thus if it could be demonstrated that the normative criteria were inappropriate, then findings from the bias/error paradigm may not be as bleak as was once thought.

Interpreting the concept of bias in terms of a social perceiver's aims, means that the use of normative criteria is inappropriate. For example, Swann (1984) contends that the accuracy of the perceiver is determined by the extent to which his or her social judgement accommodates his or her

aims within a particular social situation. His point seems valid given that a successful person in a particular environment is one who is able to precisely interpret social stimuli and act appropriately.

The point is, that normative standards set by researchers in the bias/error paradigm were unrealistic and inappropriate. This goes much of the way towards accounting for the seemingly dismal track record of the social perceiver. Thus, in the present study I decided against interpreting accuracy in some absolute sense, but rather in relation to some appropriate criteria.

Two appropriate criteria are, according to Funder (1987), behavioural prediction and inter-rater agreement. Because the aim of the selection interview is to determine which interviewees are more likely to perform well in a particular job, some measure of job performance seemed an appropriate criterion. For the job of psychology tutor, the existing measure of job performance is student evaluation and evaluation by teaching supervisors. Students complete a scale asking them to rate their particular tutor on several job dimensions. Additionally, I planned to use a similar scale to obtain an evaluation from the person directly in charge of the particular tutor, who had supervised them throughout the year.

Methodological Problems

A second issue concerns the methodological approach taken in many interpersonal perception studies. Swann (1984) contends that the reason person perception research is so unflattering is because it lacks external validity. He argues that past treatments of the social judgement process were pessimistic because the 'uniquely social aspects' of the process are

overlooked. Similarly, Cohen (1981) proposes that the experimental tasks within the laboratory simply do not resemble conditions where these judgements are used in real life. In the laboratory there is no ongoing feedback; the task often has a time limit and there is no opportunity for checking.

These factors are important. Hogarth (1981) argues that the availability of feedback and its interpretation are perhaps more meaningful than actual predictive validity. On the basis of this, a more appropriate way of studying the area would take into account the continuous and adaptive nature of the social judgement process.

Thus it would appear that judgements made in the context of the laboratory, that are classified as erroneous, may perhaps be accurate in the context of more realistic social judgement contexts. Indeed, it has been demonstrated that in certain situations people do attend to base-rate information (Ross & Fletcher, 1985).

In the present study I aimed to reflect the concerns expressed by these theoreticians. The interview situation is one with obvious external validity as it involves prediction of job behaviours on the basis of a limited interaction between the interviewer and interviewee. Thus if it can be assumed that the bias/error research conclusions regarding people's overall lack of competence are questionable, a further step is to establish peoples' actual level of accuracy.

How Accurate Are People in Their Social Judgements?

DePaulo (1978) conducted an interpersonal accuracy study in which subjects were required to predict help seeker's responses. The criterion

used was self-report of the help seekers. DePaulo found that predictive accuracy of the subjects was substantially better than chance.

Similarly, Snodgrass (1985) found that subjects were reasonably accurate in their sensitivity to the feelings and reactions of another person with whom they had been interacting. Again the criterion used was self-report of the other member of the dyad. All of the mean sensitivity scores were significantly greater than zero ($p < .001$).

Sechrest and Jackson's (1981) subjects were nurses who had known each other at least 5 months. Again the finding was that subject's predictive accuracy significantly exceeded chance levels using self-reports of the interacting members as criteria.

Bernstein and Davis (1982) used a forced choice situation to investigate the effects of perspective taking and self-consciousness on accuracy. Using self reports as a criterion they discovered that strangers' self-description predictions significantly exceeded chance.

A comparison between predictive accuracy of friends and strangers showed that both groups produced accuracy levels significantly exceeding chance, using a criterion of inter-rater agreement (Funder and Colvin, 1988). However it was also found that friends were significantly better at predicting personality traits of the ratees than were strangers.

Thus, in this study I hypothesized that subjects' predictions of job performance on the basis of a selection interview would exceed chance levels. However, I believed interviewers' interpersonal accuracy levels would be lessened as the participants were strangers to each other prior to the interview.

As has been consistently shown in the research reported above, people

can be reasonably accurate; this finding is supported across a variety of situations. Although it has been established that some level of accuracy is present, an important question to consider is *who* is more accurate.

The Influence of Acquaintanceship on Interpersonal Accuracy

A robust finding across several research areas, is that individuals acquainted with the person whose personality they are predicting, are more accurate than those individuals not acquainted with him or her. Funder and Colvin (1988) conducted a study where one group of subjects described their own personalities. These subjects were then rated by two acquaintances, as well as by two strangers. Results showed that although stranger ratings agreed with self ratings of the targets to a significant degree, friends' ratings were significantly more accurate than strangers.

It makes sense that degree of acquaintance can be a mediating influence when predicting another person's personality or future behaviour (Kenny, 1988). This is likely given the increased opportunities acquaintances have over strangers to observe the target's behaviour. Moreover, findings in the peer assessment literature support this view. Peer ratings have been shown to have substantial validity when used for predicting a diverse range of performance criteria. In Kane and Lawler's (1978) review the median validity co-efficient was 0.43.

Further, in the area of performance evaluation, it has been found that acquaintanceship with the ratee, as well as knowledge of the job, mediate evaluation accuracy (Kozlowski &Kirsch, 1987). Thus in the present study I predicted that supervisor and student evaluations (individuals who were acquainted with the interviewees), would show a significantly high

level of agreement. A further question concerns the issue of whether some interviewers are likely to be more accurate than others.

Attributional Complexity: An Individual Difference Measure

An alternative question to how accurate people in general are, concerns the issue of whether some people are more accurate in their social perceptions than others. There is evidence from both the interpersonal perception literature and selection research, that individual differences may mediate the relationship between predictor and criterion.

Bernstein and Davis (1982) examined the construct of perspective taking and found that it mediated both stereotype and differential accuracy. People high in perspective taking (as opposed to low perspective takers) adopt an attributional perspective that is more similar to the target's point of view.

Further, interviewers who were high in cognitive complexity were more likely to perceive and evaluate similarity/dissimilarity in interviewees (Leonard, 1976). Cognitive complexity represents the tendency of some people to use a widely differentiated variety of concepts to describe stimuli. Cognitively simple people are thought to use fewer concepts. Thus, I believed that it would be useful to include an individual difference measure to investigate its impact on accuracy of the interviewers.

The Attributional Complexity Scale was developed by Fletcher, Danilovics, Fernandez, Peterson and Reeder (1986). The underlying premise of the scale is that some people possess more complex attributional schemata than others. Attributionally complex people are

thought to have a higher level of intrinsic motivation to explain human behaviour, a greater awareness of the extent to which people's behaviour is a function of both internal dispositions and external causes, to indulge in meta cognitive attributional thinking, and so forth.

The scale has been used with some success to understand better the nature of interpersonal accuracy. Fletcher, Grigg, and Bull (1988) reported that attributionally complex people produced more accurate personality impressions than did attributionally simple subjects, on the basis of a fifteen minute conversation with a stranger. However, this effect was present only in the condition where subjects were motivated to develop a personality appraisal of their partner as opposed to the simple aim of having a conversation.

Further, Fletcher, Bull and Reeder (1988) randomly assigned subjects to conditions in which they wrote an essay either attacking or defending the view that homosexuality should be illegal. Other subjects were required to read these essays and to estimate the true underlying attitude of the essay writer (given that the essay writer had no choice as to the view they were to argue for). Results demonstrated that attributionally complex subjects were significantly more accurate that attributionally simple subjects when estimating the essay writers' actual attitudes.

Based on these findings I predicted in this study that attributionally complex interviewers would more accurately predict the job performance of the interviewees than would attributionally simple interviewers.

Validities Reported in Selection Interview Research

The earliest review of the selection interview literature revealed disappointingly low validity co-efficients (Wagner, 1949). Mostly, this involved estimation of a global validity co-efficient between a predictor and some relevant criterion. Of the 22 validity studies reviewed, the median correlation was r = 0.27, with co-efficients ranging from 0.09 to 0.94. The disappointing indications from the research prompted Wagner (1949) to suggest that although certain limited kinds of information can be obtained from the interview, a combination of other data gathering mechanisms interpreted in a statistical as opposed to a clinical manner, would have far greater utility.

By 1964, Mayfield noted that the literature still indicated relatively low reliabilities and validities. Encouragingly however, the most recent review in the selection interview area (Arvey and Campion, 1982), concluded that recent research findings had not been as pessimistic about the validity and reliability of the selection interview. For example, Landy (1976) conducted a field study in which the predictive validity of decisions made by a three man selection board screening police candidates were explored. The study produced low to moderate validities when compared against subsequent job performance ratings. Further, Anstey (1977) came up with a validity co-efficient of 0.66 (when corrected for restriction of range). This was based on a followup of a British Civil Service Selection Board using Civil Service ranks obtained 30 years later as a criterion.

Yet information about *why* an interviewer's decision is valid or invalid is still extremely limited (Dunnette and Borman, 1977).

Knowledge in the area could be enhanced using findings from interpersonal research. This would mean that the question of why interview decisions are valid or invalid could be better understood. Almost certainly to the area's detriment, researchers have neglected associated research in the person perception field.

> "It is as if industrial and organizational psychologists have studied the employment interview in isolation from the rest of psychology, perhaps even ignoring the fact that the phenomenon under investigation is essentially a perceptual process." (Arvey and Campion, 1982 p. 312).

The other main body of research in the selection interview area has been more microanalytical in nature. Based on the assumption that bias and error impact negatively on validity, this research has flourished. Yet despite the abundance of such studies, the nature of the relationship between interviewer bias and interviewer accuracy remains largely unexamined.

Bias and the Selection Interview

The largely unsatisfactory outcome of earlier validity research led prominent reviewers of the literature to urge a shift in focus. Mayfield (1964) recommended a reorientation toward decision-making processes inherent in the interview, the aim being to determine which factors were producing or influencing interviewer judgements. Part of this change involved utilizing a microanalytical approach as opposed to the macroanalytical emphasis of the validity research.

Consequently an abundance of studies appeared in major journals during the late 1960's and the 1970's investigating these processes. However, research focusing on process has not increased knowledge in the area in the way that Mayfield may have hoped. Despite an abundance of such research, studies investigating the linkage with the actual outcome of the interview are virtually non-existent. Yet, the pervasive assumption of much of the research is that bias is a distorting influence and that research of this type may aid in its prevention. For example, one reviewer (Wright, 1969), advocated the use of structured or patterned interviews. It was thought that this type of interview seemed to control some of the most distorting influences, such as interviewer bias, that impinge on the use of the selection interview.

In the present study I aimed to address this deficit by examining the influence of certain biases on the accuracy of the interviewer's decision. I discuss these in the following section.

Methodology

The psychology of social judgement was dominated in early years by a host of studies concerned with the accuracy of the lay perceiver. However, a stinging series of methodological critiques in the mid 1950's brought this research to a virtual halt. The effect of this was a reorientation towards investigating judgemental bias and error.

The basis of Cronbach's (1955) and Gage and Cronbach's (1955) attack was the use of discrepancy scores (as a measure of judgemental accuracy), computed simply by measuring the difference between predictor and criterion. These scores are problematic in that they represent not only

judgemental accuracy, but a host of rater biases and statistical artifacts as well. Thus, studies which had measured accuracy as a global discrepancy score were essentially uninterpretable.

Cronbach distinguished between four components of judgemental accuracy. These were: elevation, differential elevation, stereotype accuracy and differential accuracy. Differential elevation is a measurement of the ability to predict the ranking of persons across traits or situations, whereas differential accuracy measures the ability to predict differences among persons in different situations. Elevation describes a generalized bias to rate a target either favourably or unfavourably across a set of traits. Stereotype accuracy describes the ability to predict a certain trait as being relatively common or uncommon among a group of people.

The scattering of studies that have measured judgemental accuracy since the Cronbach critiques have either utilized research designs that control for variables that concerned Cronbach (Bernstein & Davis, 1982; Sechrest & Jackson, 1961), or that account for each component (Kenny & Albright, 1987; DePaulo, Kenny, Hoover, Webb, & Oliver, 1987; Harackiewicz & DePaulo, 1982; and Funder & Colvin, 1988).

Specific parts of the methodology in the present study control for, or measure these response set biases. Elevation will be controlled for by utilizing correlational measures as opposed to discrepancy scores, as advised by Cronbach (1955). This will eliminate the possibility of merely measuring some systematic response bias in the rater's judgements.

Second, stereotype accuracy will be controlled for by measuring it directly. This will be achieved by asking interviewers two weeks prior to the actual interviews to rate the average tutor's performance on a scale

measuring job performance. This will enable me to determine the extent that interviewers use stereotyping in their predictions of the supervisor and student evaluations, and whether stereotypical accuracy has an impact on the level of agreement.

Projection

A further problem in judgemental accuracy concerns projection. Predicting a certain behaviour accurately could be the result of the rater being judgementally accurate, or alternatively the rater could be assuming that the ratee is similar to him or herself, and rating them on the basis of this (Gage & Cronbach, 1955). The existence of projection has been noted in the selection interview. In an attempt to account statistically for decision variance in the interview, Sydiaha (1962) discovered that assumed similarity does occur and that this can lead to inconsistency among interviewers. Thus, I felt it important in this study to measure projection by determining the extent to which interviewers' self evaluations influenced their predictions.

Nonetheless, the relationship between projection and accuracy is not clear cut. The standard assumption is that too much reliance is placed on our own attitudes and behaviour when predicting; this is thought to be normatively incorrect. Yet, Hoch (1987) believes that projection can be functional. He contends that if people are actually similar to the rater, then it may be entirely appropriate to project. Moreover, in a predictive accuracy task he demonstrated that 65% of subjects could have enhanced their accuracy had they projected more. Hoch explained these results by arguing that subjects in this task had a hard time identifying and

weighting cues indicative of their targets, other than those contained in their own positions.

Projection in this study will be measured by having the interviewers complete a self evaluation scale two weeks before the interview, measuring their own tutoring abilities. My prediction is that increased projection by the interviewers will be positively correlated with their accuracy levels. This is because the interviewees are probably similar to the interviewers. Both groups of subjects are females and have been educated to a post-graduate level. Further, both interviewers and interviewees are employed in tutoring positions.

Summary

The major aim of this study was to investigate the selection interview from an interpersonal perspective. Further, I planned to examine predictive accuracy and judgemental bias within a realistic and externally valid context. The main hypotheses were:-

a) Interviewers would achieve levels of accuracy beyond chance levels. This corresponds with studies done so far which have found above chance levels of accuracy among social perceivers.

b) Agreement between supervisor and student evaluation would be significant and substantial. This is consistent with Funder and Colvin's (1988) comparison between the accuracy achieved by friends and by strangers.

c) As suggested by Gage and Cronbach (1955), interviewers' self evaluations would be positively related to their predictions. Further, I hypothesized that the more subjects projected, the more accurate they would be (due to the fact that interviewers resemble interviewees in several key aspects).

d) I predicted that attributionally complex interviewers would achieve greater predictive accuracy than attributionally simple people (based on research demonstrating that Attributional Complexity mediates the Fundamental Attributional Error).

METHOD

Subjects

All subjects were post-graduate students at the University of Canterbury. Interviewees were 10 female students currently employed as Stage 1 psychology laboratory tutors. Of the ten approached, all agreed to participate in the study. Interviewers were 20 female Teaching Fellows employed in faculties other than psychology. The criterion for selection was that they were to have at least one year's tutoring experience. Ages ranged from 22 to 57 years and tutoring experience from 1 to 28 years. Of the 26 people approached, 21 agreed to meet with the experimenter. Of these one refused to participate further due to time constraints.

Overview

In this study each interviewer conducted two interviews. Each interviewee was interviewed four times, making a sample size of forty interviews. Interviewers completed predictor scales requiring them to estimate the job performance of the individuals they had interviewed. Job performance measures were obtained from supervisor ratings as well as student evaluations of tutors performance. Attributional complexity ratings were also obtained to determine whether individual differences in interpersonal accuracy are present.

Data were analyzed using both within-subject (interviewers and interviewees) and across scale items for measurement of interpersonal accuracy, projection, and stereotype accuracy. Individual item analyses were also conducted to determine whether interviewers were in agreement about the same interviewee, as well as to see if certain interviewees were rated more easily than others.

Table 1

Table Demonstrating Design of Study

	Interviewees				
Interviewers	1	2	3	4	5
1	X	X			
2		X	X		
3			X	X	
4				X	X

Procedure

a) Interviewers

Interviewer subjects were approached individually and informed that this was a study examining the processes involved in selection interviews. They were informed that they would be required to conduct two separate interviews with other students. Interviewer subjects were told to try and picture themselves in a role-play situation where they were the Head Teaching Fellow interviewing prospective job applicants for the position of Stage 1 psychology tutor. Reassurance was given that they would need absolutely no knowledge of the psychological content relevant to the position, but that they were merely trying to assess tutoring and communication potential. Interviewers were informed that the aim of the interview was to predict specific job behaviours of the interviewee, on a questionnaire which they would be given immediately following the interview.

Interviewer subjects were then given two questionnaires to complete. One was a self evaluation requiring them to rate their own abilities as a tutor, and the other was a stereotype evaluation, asking them to estimate the tutoring abilities of the average tutor. Order of presentation was controlled, with half the subjects filling out the self evaluation first and the other half initially filling out the stereotype evaluation. Written instructions were also attached, as well as items asking them to provide their age, number of years spent tutoring and department (see Appendix 1).

At this time interviewers were also issued a Job Description (Appendix 2) outlining the responsibilities, person specification requirements and situations they would have to be able to deal with in the job. Interviewer subjects were asked to keep the Job Description as reference, but to return the two questionnaires via the Internal Mail System. A 100% response rate was achieved.

Interviewers were then informed that the experimenter would contact them in approximately two weeks time to arrange interview times. Meanwhile they were asked to think about questions they might ask in the coming interview.

b) Interviewees

Interviewees were approached and informed they would be required to attend four interviews. They were told to role play as if they were applying for the position they already held. At the time of the interviews, interviewees had received no formal course feedback as to their performance as tutors. Thus it was assumed that knowledge of their actual job performance would not influence their responses in the interviews. Interviewee subjects were informed that they would be interviewed four times, each time by a Teaching Fellow from another faculty.

Both interviewers and interviewees were given the experimenter's telephone number and were encouraged to approach her with any queries they might have. Additionally, they were assured that results of any questionnaires and all information conveyed in interviews would

remain confidential. Further, interviewee subjects were informed that the outcome of these interviews would in no way impact upon the position they already held.

Interviewers and interviewees were provided with written instructions (see Appendix 3) to reinforce the verbal ones at the initial meeting as it was felt that both groups should have time to prepare themselves for the interview, as they would in reality.

Subjects were then assigned identification numbers, and interview order as well as partner assignment was completely randomized. A time lapse of at least two weeks in between initial meeting with interviewers (and completion of questionnaires) and actual interviews was left. This was believed to be sufficient time to prevent memory of responses to the initial questionnaires interfering with responses to subsequent questionnaires.

c) Interview Setting

Interviews were conducted in a room containing a table and two chairs. Interviewers were free to arrange the chairs as they felt most comfortable. All subjects were asked to arrive five minutes prior to the start of the interview in order that any last minute queries could be answered and a verbal reiteration of the procedure could be given.

Interviewers were led into the room first and asked to make themselves comfortable. Next, interviewees were brought in and introduced to the interviewer. At this stage it was established whether or not the two had met previously (none had). Subjects were informed that

the interview was to last fifteen minutes and that the experimenter would return at the end of this time. The experimenter then left and the interview commenced.

At the end of fifteen minutes the experimenter knocked and entered the interview room. Next, the interviewer was given the job performance prediction form to complete (described below). Subjects were again reassured that interviewees would have no access to the forms or receive any feedback whatsoever form the interview. An identical procedure to that of the first interview would follow.

Upon completion of both interviews and questionnaires, subjects were given the Attributional Complexity Scale to complete. Finally they were thoroughly debriefed and informed as to the specific aims of the experiment. Subjects were also informed that a summary report of the findings would be forwarded to them. Interviewees were debriefed in the same manner following the last of their four interviews.

Student evaluations were obtained half-way through the year during their normal laboratory period. Supervisor ratings were obtained from the Teaching Fellows in charge of the tutors. These people had observed the tutors in laboratories during the year and were believed to have ample experience with which to provide a criterion rating, Ratings were obtained from the supervisor before they had an opportunity to view data collected from the student evaluations, in an effort to avoid confounding of the two sets of ratings.

Measures

Scale on Which all Questionnaires Were Based

All scales were based directly on an existing 25 item questionnaire currently used in course and tutor feedback (see Table 2). It is administered every year to Stage 1 psychology students and asks them to rate their own tutor on items such as: "How would you describe your tutor's attitude to students in the course?" and "The tutor is able to answer questions clearly and concisely". The questionnaire is used as feedback to the tutors and aims to guide them as to their specific strengths as well as areas to improve upon.

Scale items were developed by the Education and Advisory Research Unit (at the University of Canterbury), and were originally in the form of a 5 point Likert Scale. A catalogue of potential items are forwarded to the Head Teaching Fellow who selects items appropriate to the Stage 1 laboratory course content and format. Items remain relatively constant from year to year with minor modifications. In an attempt to increase the variance of responses, the 5 point scale was changed to a 7 point scale at the request of the experimenter.

Within this study each scale was modified slightly to remain consistent with the aim of the questionnaire. This meant that the instructions changed each time as well as the form of the items. For example on the Self Evaluation a particular item reads "Do you think you stimulate interest in the subject?". This would be rephrased to read: "The

average tutor would stimulate interest in the subject", on the Stereotype Evaluation.

Additionally, several items were removed from the modified scales (Items 11, 21 and 23) (see Table 2). These were items which involved rating the laboratory course itself and not the performance of the tutor.

Table 2

Items From Scale Used as Feedback to Stage 1 Psychology Tutors

Item Number	
1.	Rate the contribution of the tutor to this course:
2.	How would you characterize the tutor's ability to explain?
3.	Did the tutor seem well prepared for classes?
4.	The tutor's knowledge of course topics appeared to be:
5.	The tutor presented material at a level which was:
6.	The tutor was able to explain difficult material to my satisfaction:
7.	The tutor was able to answer questions clearly and concisely:
8.	The tutor stimulated my interest in the subject:
9.	Did the tutor evaluate your work in a constructive and conscientious manner?
10.	The tutor encouraged students to develop their own ideas and approaches to problems:
11.	Was class discussion a valuable part of this course?
12.	The tutor initiated fruitful and relevant discussions:
13.	Was a good balance of student participation and tutor contribution achieved?
14.	Did the tutor try to involve all students in class activities?
15.	How often was discussion monopolized by only one or a few students?
16.	How often did the tutor encourage interaction among students?
17.	Was the tutor receptive to differing viewpoints or opinions?
18.	How would you describe the tutor's attitude toward students in
the	course?
19.	How patient was the tutor in working with you?
20.	How helpful was the tutor to students with problems?
21.	Have the labs seemed a valuable part of the psychology course?
22.	What the tutor expected of students was:
23.	Did you feel class time was spent on unimportant and irrelevant material?
24.	How accessible was the tutor to students after class?
25.	The tutor was sensitive to student needs and concerns:

Self Evaluation

This scale was completed by the interviewers at least two weeks prior to the interviews (see Appendix 4). General instructions were: "Based on your performance as a Teaching Fellow how well would you rate yourself on the following items?".

Stereotype Evaluation

This questionnaire was also filled out by the interviewers at least two weeks prior to the interviews (see Appendix 5). General instructions were: "Based on your experience as a Teaching Fellow how do you feel the average tutor (in any department) would perform on the following items?"

Interviewer Prediction of Interviewee's Job Performance

The predictor was completed by the interviewer immediately following each interview (see Appendix 6). Instructions were: "Based on the interview you have just given how do you think the average hypothetical Stage 1 Psychology student would rate the person you have just interviewed on the following items?"

Criterion Measures

These scales were also based on the same questionnaire as those above.

a) Supervisor Evaluation This scale was completed by the Psychology Teaching Fellow responsible for the interviewee during the year (see Appendix 7). Teaching Fellows supervise two tutors in each laboratory class. Instructions were: "Based on your observation of the tutor's performance this year how would you rate her on the following items."

b) Student Evaluation Stage 1 psychology students were asked to rate their laboratory tutors (see Appendix 8), and were given the following instructions: "Based on your observation of the tutor's performance this year how would you rate her on the following items?"

Attributional Complexity Scale

The Attributional Complexity Scale is an individual differences questionnaire developed by Fletcher et al (1986) (see Appendix 9).

RESULTS

Recall that each interviewer conducted two separate interviews and each interviewee was interviewed by four different interviewers. The aim of the interviews was for the interviewers to predict job performance of the interviewees. The interviewees were in reality already employed in the position. This meant that actual job performance could be measured. Ratings of the tutors were obtained from both their supervisors and the students whom they tutored. Additionally, interviewers completed scales measuring evaluations of their own tutoring abilities, as well as their ratings of the 'average hypothetical' tutor. Interviewers also completed an Attributional Complexity Scale.

The results will be presented in five major sections. First, I will examine whether there is a significant overall level of agreement between interviewers who interviewed the same interviewee. Second, I will look at the agreement between supervisors' and students' ratings of the same tutor. Third, I will examine whether interviewers achieved a significant level of accuracy. Fourth, I will look at the influence of projection and stereotyping on overall accuracy. Finally, I will investigate other possible influences on the level of accuracy.

Accuracy of Interviewers

Overall Level of Agreement Between Interviewers

Each interviewee was rated by four different interviewers. To determine whether there was significant agreement between the interviewers, intra-class correlations were calculated between each item on the four interviewer predictions of job performance of the interviewees (using each interviewee as a grouping variable). Intra-class correlations test for non-independence in nested designs (see Kenny & LaVoie, 1985, for more details).

Each scale had 22 items. Except for two items, all 22 correlations were low and close to zero (all correlations were non-significant). Interviewers exhibited little agreement concerning the interviewees that they had jointly interviewed. On the basis of these findings, data from each interview were treated as independent.

Overall Agreement Between Interviewer Predictions and Supervisor and Student Ratings

Interviewers completed estimations of interviewees' job performance, based on their observations in the interviews. Each interviewer's set of prediction ratings were correlated with the ratings of the supervisor and students, of the same interviewee within-subject and across items in each scale. These within interviewer subject correlations (converted to Fisher's z scores), were used as measures of accuracy for the interviewers'

predictive judgements.

If the number of significant accuracy correlations was due to chance effects only, approximately four out of 40 (total number of interviews conducted) correlations, between interviewer predictions and supervisor and student evaluations would be significant at the $p < .05$ level (five per cent of the total).

As can be seen in Table 3, the number exceeded chance levels. Using supervisor evaluations as criteria, there were a total of fifteen positive significant ($p < .05$) accuracy correlations. Using student evaluations, a total of thirteen co-efficients were significant ($p < .05$). Thus my predictions was supported. This is an impressive finding, given the inherent difficulty of the task.

Table 3

Number of Accuracy Correlations (significant at the $p < .05$ level) Between Interviewer Predictions of Job Performance of Interviewee, and Supervisor and Student Evaluations

Criterion	Interview 1	Interview 2
Supervisor Evaluation	8	9
Student Evaluation	7	6

Overall accuracy of the interviewers was measured in an alternative way. The within interviewer subject and across item correlations between predictions of job performance (made by the interviewers) and ratings of job performance (made by the interviewees' supervisors and students) were converted to Fisher's z scores. They were then averaged across interviewer separately for correlations between supervisor and student evaluation. Although positive, neither of these averaged co-efficients (converted back to correlations from Fisher's z scores) were significant (see Table 4).

Table 4

Averaged Within Interviewer Correlations Between Interviewer Predictions of Job Performance of Interviewee, and Supervisor and Student Evaluations

Interviewer Predictions	Supervisor Evaluation	Student Evaluation
Interview 1	.29	.33
Level of Significance	ns	ns
Interview 2	.27	.29
Level of Significance	ns	ns

Overall Level of Agreement Between Supervisor and Student Evaluations

Interviewees were rated by their supervisors as well as by the students they were tutoring during the year. Correlations were calculated across each set of items and within-subjects, on the supervisor and student evaluations scales for each interviewee. Data was analyzed separately for the first and second interviews.

These correlations were converted to Fisher's z scores and a mean z score between the supervisor and student evaluations was calculated and converted back to a correlation. Results showed a reasonable level of agreement existed between the supervisor ratings and the student evaluations. For the first interview, the mean correlation was 0.39 ($p < .05$), and for the second interview was 0.44 ($p < .05$).

Consistency of Interviewers Between Predictions Made on First and Second Interviews

Prior to analyzing the Attributional Complexity data, it was necessary to establish whether interviewers' level of accuracy was consistent across the two interviews. Without this consistency, it is not possible for an individual difference measure to predict behaviour. Correlations were calculated between the interviewers' predictions from their first interview, and both supervisor and the student evaluations (across items and within-subject). The same procedure was repeated using the predictions from the second interviews. These two sets of accuracy

correlations (between the two interviews), were converted to Fisher's z scores. An overall correlation was then calculated across subjects between accuracy co-efficients on the first and second interviews. Using supervisor evaluations as criteria and converting z scores back to correlations, subjects who were accurate on their first interview were not necessarily accurate on their second interview(r = -.2, ns). Identical results were found using the student evaluations of job performance (r = -.07, ns). Both correlations were negative and close to zero. As interviewers were not consistently accurate, the Attributional Complexity data was not analyzed.

Projection

Was There Any Evidence of Interviewers Projecting?

Interviewers completed a scale, rating themselves as tutors. Projection was defined as the extent to which interviewees' self evaluations agreed with their predictions of the interviewees' job performance. This was measured within-subject and across items in the scale. Separate correlations were calculated for their first and second interview. These correlations were then converted to Fisher's z scores and averaged across each interviewer. .Although in the hypothesized direction, neither average correlation was significant. The mean correlation for the first interview was .304 (ns), and for the second interview was .30 (ns). This showed that interviewers were not projecting to a significant extent.

Did Projection Influence the Predictive Accuracy of the Interviewers?

Although the results showed that subjects were not projecting to a great extent, I was interested in assessing what happens to accuracy levels when the level of projection varies. As previously described, for each interviewer I obtained a measure of projection by calculating a within-subject and across item correlation between each interviewer's self evaluation and their predictions of job performance. Second, each interviewer's actual accuracy was calculated by correlating predictions of interviewees' job performance with both the supervisor and student evaluations. These two sets of correlations were converted to Fisher's z scores. Finally, the projection z scores were correlated with the accuracy z scores and reconverted back to correlations. A positive correlation (reconverted from Fisher's z scores), would indicate that a high level of projection would be associated with a high level of predictive accuracy. Data was again analyzed independently from the first and second interviews.

For the first interviews projection was positively and significantly related to accuracy ($p < .05$) (see Table 5) using both supervisor and student evaluations. However, this result was not replicated in the second interviews. Both correlations here were non-significant. The results from the first interviews do however support my hypothesis that higher projection would increase the accuracy levels of the interviewers.

Table 5

Overall Correlations Representing the Relationship Between Degree of Projection and Predictive Accuracy of Interviewers

	Accuracy Correlations			
	Student		Supervisor	
	Criterion	Significance	Criterion	Significance
Projection Interview 1	.51	.05	.42	.05
Projection Interview 2	-.14	ns	.15	.05

Stereotype Accuracy

Stereotype accuracy is a hypothesized component of overall accuracy as proposed by Cronbach (1955). Cronbach argues that the extent to which an individual is accurate may depend on the extent to which a particular target is similar to the rater's stereotype of a particular group. Stereotype accuracy was measured by asking interviewers to rate how they felt the average tutor would be rated on a measure of job performance. This procedure enabled me to control for the interviewers' stereotypical accuracy, as opposed to their differential skill in accurately rating a particular individual.

Identical procedures to those used to analyze the projection data were utilized. The extent to which interviewers used their stereotypes to aid in their prediction of the interviewees' job performance was measured by correlating the interviewers' stereotype evaluations with their predictions of job performance (across scale items and within-subject). Again, independent analyses were conducted across the first and second interviews. None of these correlations were significant. Thus interviewers did not appear to be using their stereotypes of the average tutor in making their predictions.

These correlations were used to determine whether stereotyping was related to interviewers' predictive accuracy levels. The correlations between stereotype evaluation and prediction of job performance were correlated with the accuracy correlations (between predictions of job performance and supervisor and student ratings of job performance). Once again these correlations were low and close to zero.

Alternative Influences on Accuracy

Were Certain Tutors Being Rated More Easily than Others?

A further explanation of these results rests with the interviewees themselves. If interviewers are not being consistently accurate or inaccurate across their interviews, then it is plausible that certain interviewees were more easily rated than others.

Intra-class correlations were used to test this possibility. Individual analyses were conducted for the two measures of job performance (supervisor and student evaluation accuracy co-efficients). Interviewees were interviewed four times, meaning there were a total of four accuracy correlations for each individual. If there was a positive and significant level of agreement between these four correlations, this would suggest that accuracy is more a function of the interviewee than the interviewer. However, results showed that this was not the case. Intra-class correlations were calculated across items and within interviewee; none were significant.

Was Age and Tutoring Experience of Interviewers a Determinative Factor?

If interviewers are relatively inexperienced as tutors and are younger, they may be less likely to know exactly what to look for in another tutor. To test this possibility, age was correlated with both accuracy correlations (supervisor and student evaluations). Identical analyses were conducted for the number of years tutoring. All correlations were however non-

significant and close to zero.

Summary

The results of this study show that a significant number of the interviewers were accurate in predicting the job performance measures of the individuals they interviewed. However the average accuracy correlations across interviewers were not significant, although in the hypothesized positive direction. Further, interviewers who interviewed the same tutor, were not in agreement in their predictions of job performance. Individuals who were familiar with the interviewee (supervisors and students) were in agreement.

Interviewers were not consistently accurate or inaccurate across the two interviews they conducted. This meant that the individual difference data was not analyzed. Overall, interviewers did not appear to be using evaluations of their own tutoring abilities (self evaluations) to any great extent. However when projection was used, there was partial evidence to suggest that accuracy levels were enhanced.

Interviewers were not stereotypically accurate and in fact made minimal use of their ratings of the average tutor. Certain tutors were not rated more accurately than others. Further, the age and tutoring experience of the interviewer was not related to interviewer accuracy levels.

DISCUSSION

The major aim of this study was to investigate how accurately individuals could predict specific job behaviours of interviewees on the basis of a short interview. The study also examined factors which might influence accuracy.

The results showed that interviewers in general did not achieve a significant level of accuracy. Interviewers conducted two interviews and in the first interview those individuals who used evaluations of their own tutoring abilities in their predictions were more accurate than those who didn't. However, this result was not replicated in the second interview. Interviewers were not consistently accurate or inaccurate across interviews. Therefore the individual difference scale measuring attributional complexity was not analyzed. Individuals who were familiar with the interviewee (their supervisors and the students they had tutored), had a significantly high level of agreement. These results have important implications for both the selection interview as well as interpersonal accuracy research. These will be discussed in four major sections.

General Level of Accuracy in Interviewer Predictions of Job Performance

I hypothesized that the predictive accuracy achieved by the interviewers would significantly exceed chance levels. This prediction was not supported. Although positive, the average accuracy correlations

of the interviewers were not significant. This finding contrasts with a scattering of similar studies investigating the predictive accuracy of the social perceiver (see DePaulo, 1987; Snodgrass, 1985; Sechrest & Jackson, 1981; Bernstein & Davis, 1982; and Funder & Colvin, 1988). This raises the question of why selection interviewers are not as accurate as raters in other interpersonal accuracy studies.

Part of the explanation for the low level of the averaged correlations between interviewer predictions and supervisor and student ratings lies in the nature of the judgements themselves. It is clear from earlier research that certain types of judgements are made more easily than others. For example, several studies have shown that people are more accurate at rating traits high in subjective visibility, such as extraversion (Funder, 1980a; Funder & Dobroth, 1987; and Funder & Colvin, 1988).

This result has been replicated in the selection interview research. Validity co-efficients are typically higher and more likely to be significant for those dimensions of behaviour more easily observed (Arvey and Campion, 1982). Similarly, Wagner's (1949) review concluded that validity and reliability were acceptable for one area, that of sociability. Further, Otis, Campbell and Prien (1962) found that psychologists were able to predict personal relations to a significant extent on the basis of an interview.

The judgements the interviewers made in this study were more subtle and complex than judgements of personality. Interviewers were required to predict how a particular interviewer would behave in a certain context. This is an inherently difficult task. Harackiewicz and DePaulo (1982) showed that although subjects were good at ranking

persons on a particular trait and ranking the liklihood of certain situations across people, they were not good at predicting interaction effects between person and situation. Further, as interviewers were required to predict specific job behaviours on the basis of a very short period of observation, high levels of accuracy should not be expected (Kenrick & Funder, 1988).

Moreover, these interviewers had little or no experience at the task. In all but one case, none of them had had experience at interviewing. These explanations go some way to explaining the low levels of predictive accuracy. However, the finding that the number of significant accuracy correlations exceeded chance levels, shows that some degree of accuracy must have been present in the interviewers' predictions. This is admirable given the difficulty of the task.

These results can be explained further using results from structured interviewing. In a structured interview, a pre-determined set of questions is asked of all interviewees applying for a particular position (Pursell, Campion & Gaylord, 1980). By using this technique it appears that interviewers can achieve substantial levels of validity. For example, a study conducted by Campion, Pursell and Brown (1988) obtained a corrected predictive validity co-efficient of 0.56. This result is consistent with other selection interview studies using the structured technique (see Pursell, Campion & Gaylord, 1980).

The success of this method is attributed to the extreme structuring and standardization of the interviews. All questions were based on a job analysis, the same questions were asked of each candidate, and anchored rating scales were used for scoring answers.

The structuring technique was not included in the present study, despite the encouraging validities found in most of the above research. This is because I wanted the interviewers to implement their own strategies and interviewing methods. As I was measuring interpersonal accuracy, the existence of a predetermined interview structure would be highly likely to have a confounding influence on the outcome of the study. Accordingly, I would expect that had structured interviews been used, the validity correlations would have been more substantial.

Projection and Interpersonal Accuracy

As predicted, when interviewer's self evaluations correlated highly with their predictions of the interviewee's job performance, predictions were more accurate. This effect however was only present in the first of two interviews conducted.

This finding supports Hoch's (1987) contention, that where rater and ratee are actually similar to each other, then projection is actually a useful strategy. As argued previously, the interviewer and interviewee were similar to each other in important respects.

These results could be explained in the following way. If the projection bias was used in conditions of uncertainty (unfamiliar situation, low job knowledge and no prior knowledge of the interviewee), I would predict that the level of bias would not necessarily lessen the accuracy of the perceiver. Indeed in this situation it may be quite useful to utilize this type of strategy.

This finding lessens the implications this study may have had for the selection interview. Normally, interviewers in more realistic situations

would not have such low job and task knowledge. However, in most cases they would still be unfamiliar with the interviewee.

Kenny (1988) proposes an important point concerning the interpretation of the projection data; namely, is a high level of agreement between an interviewer's self evaluation and predictions of job performance a result of assumed similarity, or is it merely a generalized response bias affecting both sets of ratings in the same way? The results from this study suggest that the former position is more appropriate for my data; although interviewer subjects weren't projecting very much, their accuracy levels were still influenced to a significant degree (in the first interview).

As was discussed previously, the implication from the bias/error process paradigm is that biases negatively influence the accuracy of social perceivers. Yet the results from the projection data contradict this, and provide support for definitions that bestow a functional quality to the notion of bias.

Stereotype Accuracy

Stereotype accuracy was measured so it could be controlled for (according to Cronbach, 1955). However as can be seen from the results, interviewers' stereotypes of the average tutor did not affect their predictions of job performance of the interviewee, nor their actual accuracy.

One possible explanation for this is that interviewers simply had no overall stereotype of the average tutor. Indeed several interviewers mentioned the difficulty they had with this task. The reason for the

problem is twofold. First, tutoring styles and methods differ from department to department within the university. Second, many of the interviewers had little knowledge of other tutors outside of their own department.

Thus the stereotype evaluation may not have been measuring stereotype accuracy in the way it was intended. Further, it is obvious from these results stereotypes were not influencing interviewer predictions of the interviewees' job performance. It is possible this occurred because interviewers failed to form an adequate stereotype.

Individual Differences

As was explained previously, the Attributional Complexity Scale was not analyzed. The rationale for this was that interviewers were not consistently accurate or inaccurate across interviews. It is difficult to explain why the interviewers were so inconsistent. However attributional complexity may still have been a determinative factor. The reason interviewers were not consistently accurate may have been because their strategies were changing from one interview to another, and consequently their accuracy may have varied.

A further explanation of the lack of consistency of the interviewers stems from ideas proposed by Kenny and Albright (1987). These theoreticians argue that most research ignores the interactive nature of social perception. They contend that in this type of situation, the ratee can actually change when interacting with different perceivers.

This type of explanation is consistent with results from the present study, including the finding that interviewers were not consistently

accurate across interviews. Further, some interviewees were not rated more accurately than others. This would suggest that accuracy is perhaps a function of the interaction, rather than of any particular feature of the interviewer or interviewee. The results from Heneman, Schwabb, Huett, and Ford's (1979) study support this view. This study examined validity of interviewer predictions of job performance. Results showed that the largest source of variance was the Judge x Interviewee interaction.

It has been shown previously that attributional complexity only influences accuracy under conditions where subjects are highly motivated (Fletcher, Grigg & Bull, 1988). Future research could also look to manipulating motivation as well as investigating the situations in which attributional complexity mediates interpersonal accuracy.

Agreement Between Supervisor and Student Evaluations

As predicted, the correlation between the supervisors' and the students' evaluations of the interviewees were positive and significant. This is consistent with the findings from Funder and Colvin's (1988) friend/stranger study, as well as results from peer assessment and performance evaluation research. The common finding here is that acquaintances can be accurate raters of variables such as the personality of the person they are rating, and their future job performance.

One explanation as to why the judgemental abilities of acquaintances are so good, lies in the opportunities acquaintances have to tap data. However, it has been shown within the peer nomination research that the validity of nominations for predicting leadership performance develops very early in the life of the group, and reaches a plateau after no

more than three weeks (Hollander, 1956b). Thus it seems opportunity for behavioural observations may not provide a full explanation of these results.

An alternative explanation is that interaction is necessary for valid ratings. Imada (1982) conducted a study to determine which view is more appropriate. Ratings made after observation of group members alone were compared with ratings made after interaction between subjects. Results showed that interaction between group members was necessary to form valid ratings. Observation alone was not sufficient.

Further, Snodgrass (1985) found that status of interacting group members affects validity of predictions. Subjects were required to interact in groups, some as leaders, others as subordinates. Results clearly showed that subordinates were more sensitive than leaders to the feelings of the other group members. This would suggest that the interactive context can affect the social perceptions of perceivers and consequently their accuracy.

Conclusions and Implications

Selection Interview Area

1) Findings from this study clearly suggest that it is useful to consider selection interviewing as if it were a particular example of interpersonal perception. This would enable researchers to answer to answer such questions such as *why* selection interviews have tended to produce low to moderate validities.

2) Results from the measurement of projection bias imply that the mass of research conducted on selection interviewer biases may not have such bleak implications for the selection interviewer as was once thought. It is obvious that in this case use of bias enhanced the level of accuracy. A more useful approach would be to assess the impact of different biases on selection interview validities independently.

Interpersonal Perception Area

1) This research shows how useful it is to investigate interpersonal accuracy within a given context. Obviously, in a selection interview, people are not as accurate as they are in other social judgemental situations.

2) Interpersonal accuracy may be better understood by assuming it is an interactive process (consistent with Kenny & Albright's, (1987) Social Relations Model). According to this model, interpersonal accuracy is the result of the actor, the partner, as well as some interaction effect between two. Results from this study suggest a fruitful line of research would be to investigate when and where individuals are accurate, especially as it has now been established that people can achieve a reasonable level of interpersonal accuracy.

REFERENCES

Anstey, E. (1977). A 30-year follow-up of the CSSB procedure, with lessons for the future. *Journal of Occupational Psychology, 50,* 149-159.

Arvey, R. D., and Campion, J. E. (1982). The employment interview: A summary and review of recent research. *Personnel Psychology, 35,* 281-322.

Bernstein, W. M. & Davis, M. H. (1982). Perspective-taking, self-conciousness, and accuracy in person perception. *Basic and Applied Social Psychology, 3,* 1-19.

Campion, M. A., Pursell, E. D., & Brown, B. K. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. *Personnel Psychology, 41,* 25-42.

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *The Behavioural and Brain Sciences, 4,* 317-370.

Cronbach, L. J. (1955). Conceptual and methodological problems in interpersonal perception, *Psychological Bulletin, 52,* 177-193.

DePaulo, B. M. (1978). Accuracy in predicting situational variations in help-seekers responses. *Personality and Social Psychology Bulletin, 4,* 330-333.

DePaulo, B. M., Kenny, D. A., Hoover, C. W., Webb, W., & Oliver, P. V. (1987). Accuracy of person perception: do people know what kinds of impressions they convey? *Journal of Personality and Social Psychology, 52,* 303-315.

Dunnette, M. D., & Borman, W. C. (1977). Personnel selection and classification systems. *Annual Review of Psychology, 30,* 477-525.

Fischoff, B. (1975). Hindsight = foresight: The effect of outcome knowledge opn judgement under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288-299.

Fletcher, G., Danilovics, P., Fernandez, G., Peterson, D., & Reeder, G. D. (1986). Attributional complexity: An individual differences measure. *Journal of Personality and Social Psychology, 51,* 844-875.

Fletcher, G .J. O., Grigg, F., & Bull, V. (1988). The organization and accuracy of personality impressions: Neophytes versus experts in trait attribution. *New Zealand Journal of Psychology, 17,* 68-77.

Fletcher, G .J .O., Bull, V., & Reeder, G. D. (1988). Bias and Accuracy in Trait Attribution: The Role of Attributional Complexity. Unpublished manuscript, University of Canterbury.

Funder, D. C. (1980a). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality, 48,* 473-493.

Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgement. *Psychological Bulletin, 101,* 75-90.

Funder, D. C., & Colvin, R. (1988). Friends and strangers: Acquaintanceship, agreement and the accuracy of personality judgement. *Journal of Personality and Social Psychology, 55,* 149-158.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52,* 409-418.

Gage, N. L., & Cronbach, L. J. (1955). Conceptual and methodological problems in interpersonal perception. *Psychological Review,, 62,* 411-422.

Harackiewicz, J. M., & DePaulo, B. M. (1982). Accuracy of person perception: A component analysis according to Cronbach. *Personality and Social Psychology Bulletin,, 8,* 247-256.

Heneman, H. G., Schwab, D. P., Huett, D. L., & Ford, J. L. (1979). Interviewer validity as a function of interview structure, biographical data, and interview order. *Journal of Applied Psychology, 60,* 748-753.

Hoch, S. J. (1987). Perceived consensus and predictive accuracy: The pros and cons of projection. *Journal of Personality and Social Psychology, 53,* 221-234.

Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgemental heuristics, *Psychological Bulletin, 90,* 197-217.

Hollander, E. P. (1956b). Interpersonal exposure time as determinant of the predictive utility of peer ratings. *Psychological Reports, 2,* 445-448.

Imada, A. D. (1982). Social interaction, observation, and stereotypes as determinants of differentiation in peer ratings. *Organizational Behavior and Human Performance, 29,* 397-415.

Kahneman, D. & Tversky, A. (1973). On the psychology of prediction, *Psychological Review, 80,* 237-251.

Kane, J. S. & Lawler, E. E. (1978). Methods of peer assessment. *Psychological Bulletin, 85,* 555-586.

Kenny, D. A. (1988). Interpersonal perceptions: A social relations analysis. *Journal of Social and Personal Relationships, 5,* 247-261.

Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin, 102,* 390-402.

Kenny, D. A., & LaVoie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology, 48,* 339-348.

Kenrick, D. T., & Funder, D. C. (1988). Profitting from controversy: Lessons from the person-situation debate. *American Psychologist, 43,* 23-34.

Kozlowski, S. W. J., & Kirsch, M. P. (1987). The systematic distortion hypothesis, halo and accuracy: An individual-level analysis. *Journal of Applied Psychology, 72,* 252-261.

Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgement. *European Journal of Social Psychology, 13,* 1-44.

Landy, F. J. (1976). The validity of the interview in police officer selection. *Journal of Applied Psychology, 61,* 193-198.

Leonard, R. L. Jr. (1976). Cognitive complexity and the similarity-attraction paradigm. *Journal of Research in Personality, 10,* 83-88.

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review, 90,* 215-238.

Mayfield, E. C. (1964). The selection interview: A re-evaluation of published research. *Personnel Psychology, 17,* 239-260.

Nisbett, R. E., & Borgida, E. (1975). Attribution and the psychology of prediction. *Journal of Personality and Social Psychology, 32,* 932-943.

Otis, J. L., Campbell, J. T., & Prien, E. P. (1962). Assessment of higher-level personnel: VII. The nature of assessments. *Personnel Psychology, 15,* 441-446.

Pursell, E. D., Campion, M. A., & Gaylord, S. R. (1980). Structured interviewing: Avoiding selection problems. *Personnel Journal, 59,* 907-912.

Ross, M. & Fletcher, G. J. O. (1985). Attribution and social perception. In G. Lindzey & E. Aronson (Eds.), *Handbook of Social Psychology,* (3rd ed., Vol.2, pp. 73-122). New York: Random House.

Sechrest, L., and Jackson, D. N. (1961). Social intelligence and accuracy of interpersonal predictions. *Journal of Personality, 29,* 167-182.

Snodgrass, S. E. (1985). Women's intuition: The effect of subordinate role on interpersonal sensitivity. *Jounal of Personality and Social Psychology, 49,* 146-155.

Swann, W. B. Jr. (1984). Quest for accuracy in person perception: A matter of pragmitics. *Psychological Review, 91,* 457-477.

Sydiaha, D. (1962). Interviewer consistency in the use of empathic models in personnel selection. *Journal of Applied Psychology, 46,* 344-349.

Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 125,* 1124-1131.

Wagner, R. (1949). The employment interview: A critical summary. *Personnel Psychology, 2,* 17-46.

Wright, O. R., Jr. (1969). Summary of research on the selection interviews since 1964. *Personnel Psychology, 22,* 341-413.

Appendix 1

Attached to this sheet you will find a Job Description for the position of Stage I Psychology Tutor. Please read this carefully before moving on to answer the two questionnaires.

The questionnaires are based on actual forms which are completed by Stage 1 Psychology students during the course of the year. The data are used as part of Tutor assessment and feedback.

Number of years tutoring:_____

Department:_____

Age:_____

Appendix 2

JOB DESCRIPTION FOR THE POSITION OF STAGE I PSYCHOLOGY TUTOR

Classification Information

Immediate Superior:- Teaching Fellow.

Responsibilities

Primary Function

Responsible for teaching Psychology Stage I lab course to small groups (10 - 15) students.

Responsibilities and Duties

a) Daily- Teach and supervise the running of each week's set lab work. Appropriate handouts are obtained from the Teaching Fellow. Student attendance is recorded and lab work is collected, marked and returned to students. They need to become familiar with requirements for the next lab.

b) Irregular- Answer student's questions about aspects of lab course, psychology lectures and possibly statistics. Tutors need to refer difficult questions or problems with students to their Teaching Fellow or the senior Teaching Fellow.

External Relationships

Within the department:- Other than to the Teaching Fellow and Senior Teaching Fellow, they need to interact with secretaries, technicians, and perhaps lecturers, although most of the 'external relationships' are carried out by Teaching Fellows or the Senior Teaching Fellow.

Limits of Financial Authority

Tutors are allowed to charge the department for lab related photocopying.

Confidential Information Available to Tutor

Through the Teaching Fellow or Senior Teaching Fellow, tutors have access to student records and former student's marks.

Promotional Opportunities

Teaching Fellow
Senior Teaching Fellow

Nature of On The Job Training

A concentrated half day training course before the start of the year's lab course. Teaching Fellows make sure tutors are kept informed of future duties.

Person Specification Requirements

Formal Education

Post graduate thesis students in Psychology are preferred, but papers students are acceptable. Occasionally, keen undergraduates with special qualities are accepted.

Experience

Several years experience as a Psychology student is necessary.

Specific Knowledge and Skills

Knowledge of workings of the Psychology Department is vital for this position, as is some ability to relate to students and impart knowledge. Where possible tutors are selected with these qualities in mind.

Personal Qualities

Friendliness and helpfulness are prime qualities of a Tutor. Also important are punctuality and reliability.

Critical Incidents :- (examples of critical incidents which the Tutor may need to face and which have an important bearing on success or failure in this position).

a) Controlling a disruptive student or a noisy, inattentive class. Failure

to control this sort of problem could - in the worse possible situation - lead to abandoning the class for that day.

b) Maintaining student attention - if the Tutor doesn't or can't, other problems follow e.g. nothing is learned.

c) If the Tutor doesn't keep a careful record of all marks and attendance, this can lead to problems at the end of the year in giving grades and marks for lab work.

d) Individual based problems- e.g. a student who is constantly absent; or a student is having great difficulty mastering the material; or a student comes to the tutor with a serious personal problem.

Appendix 3

INSTRUCTIONS

Try to imagine you are the Head Teaching Fellow/Tutor for the Psychology Department and that you are interviewing prospective job applicants. Use the Job Description you have been provided with on which to base your interview questions. You are free to conduct the interview in what ever manner you wish.

The interview will last 15 minutes. After this you will be asked to complete a short questionnaire very similar to the ones you have completed previously. The aim of the questionnaire will be for you to predict the job performance of the person you have just interviewed.

Thank you very much for your co-operation.

INSTRUCTIONS

Imagine you are applying for the position you already possess, that of Stage 1 Psychology Tutor. Think about the sorts of information you would wish to convey in a job interview. The interview will last 15 minutes.

None of the content of this interview will have any bearing on your present position and the only person to have any access to any data coming directly from the interview situation will be myself.

Thank you very much for your co-operation and support.

Vivienne Bull

Appendix 4

SELF EVALUATION

For each question, please CIRCLE THE NUMBER which best reflects your opinion. Based on your performance as a Teaching Fellow how well would you rate yourself on the following items:

1. Rate your contribution to the course:	Poor 1 2 3 4 5 6 7 Excellent
2. How would you characterize your ability to explain:	Very poor 1 2 3 4 5 6 7 Excellent
3. Are you well prepared for classes:	No, never 1 2 3 4 5 6 7 Yes, always
4. How would you rate your knowledge of course topics:	Inadequate 1 2 3 4 5 6 7 More than adequate
5. Do you think you present material at a level appropriate to class:	Almost never 1 2 3 4 5 6 7 Almost always
6. Are you able to explain difficult material to the student's satisfaction:	Almost never 1 2 3 4 5 6 7 Almost always
7. Are you able to answer questions clearly and concisely:	Almost never 1 2 3 4 5 6 7 Almost always
8 Do you think you stimulate interest in the subject:	Not at all 1 2 3 4 5 6 7 Very much
9. Do you evaluate work in a constructive and conscientious manner:	Definitely no 1 2 3 4 5 6 7 Definitely
10. Do you encourage students to develop their own ideas and approaches to problems:	Never 1 2 3 4 5 6 7 Frequently
11. Do you initiate fruitful and relevant discussions:	Never 1 2 3 4 5 6 7 Frequently
12. Is a good balance of student participation and tutor contribution achieved:	No, very poor 1 2 3 4 5 6 7 Yes very good
13. Do you try to involve all students in class activities:	No, seldom 1 2 3 4 5 6 7 Yes, consistently
14. How often is discussion monopolized by only one or a few students:	Almost never 1 2 3 4 5 6 7 Almost always
15. How often do you encourage interaction among students:	Never 1 2 3 4 5 6 7 Frequently

16 Are you receptive to differing viewpoints or opinions: No, never 1 2 3 4 5 6 7 Yes always

17. How would you describe your attitude toward students in the course: Indifferent 1 2 3 4 5 6 7 Very helpful

18. How patient are you when working with students: Very impatient 1 2 3 4 5 6 7 Very patient

19. How helpful are you to students with problems: Not at all helpful 1 2 3 4 5 6 7 Very helpful

20. How clear are your expectations of students:: Very unclear 1 2 3 4 5 6 7 Very clear

21. How accessible are you to students outside class hours: Very inaccessible 1 2 3 4 5 6 7 Very accessible

22. Do you think you are sensitive to student needs and concerns: Almost never 1 2 3 4 5 6 7 Almost always

Tutor Code: _____

Appendix 5

PSYCHOLOGY TUTOR EVALUATION

For each question, please CIRCLE THE NUMBER which best reflects your opinion.. Based on your experience as a Teaching Fellow, how do you feel the average tutor (in any department) would perform on the following items?

1. Rate the contribution of the average tutor to a course: Poor 1 2 3 4 5 6 7 Excellent

2. How would you characterize the average tutor's ability to explain: Very poor 1 2 3 4 5 6 7 Excellent

3. Would the average tutor be well prepared for class: No, never 1 2 3 4 5 6 7 Yes, always

4. The average tutor's knowledge of course topics would be: Inadequate 1 2 3 4 5 6 7 More than adequate

5. The average tutor would present material at a level appropriate to class: Almost never 1 2 3 4 5 6 7 Almost always

6. The average tutor would be able to explain difficult material to students' satisfaction: Almost never 1 2 3 4 5 6 7 Almost always

7. The average tutor would be able to answer questions clearly and concisely: Almost never 1 2 3 4 5 6 7 Almost always

8 The average tutor would stimulate interest in the subject: Not at all 1 2 3 4 5 6 7 Very much

9. The average tutor would evaluate work in a constructive and conscientious manner: Definitely no 1 2 3 4 5 6 7 Definitely

10. The average tutor would encourage students to develop their own ideas and approaches to problems: Never 1 2 3 4 5 6 7 Frequently

11. The average tutor would initiate fruitful and relevant discussions: Never 1 2 3 4 5 6 7 Frequently

12. The average tutor would achieve a good balance of student participation and tutor contribution: No, very poor 1 2 3 4 5 6 7 Yes very good

13. Would the average tutor try to involve all students in class activities: No, seldom 1 2 3 4 5 6 7 Yes, consistently

14. For the average tutor, how often would the discussion be monopolized by only one or a few students: Almost never 1 2 3 4 5 6 7 Almost always

15. How often would the average tutor encourage interaction among students: Never 1 2 3 4 5 6 7 Frequently

16. The average tutor would be receptive to differing No, never 1 2 3 4 5 6 7 Yes always

viewpoints or opinions:

17. How would you describe the average tutor's attitude toward students in the course: Indifferent 1 2 3 4 5 6 7 Very helpful

18. How patient would the average tutor be in working with students: Very impatient 1 2 3 4 5 6 7 Very patient

19. How helpful would the average tutor be to students with problems: Not at all helpful 1 2 3 4 5 6 7 Very helpful

20. What the average tutor would expect of students would be: Very unclear 1 2 3 4 5 6 7 Very clear

21. How accessible would the average tutor be to students outside class hours: Very inaccessible 1 2 3 4 5 6 7 Very accessible

22. The average tutor would be sensitive to student needs and concerns: Almost never 1 2 3 4 5 6 7 Almost always

Tutor Code: _____

Appendix 6

PSYCHOLOGY TUTOR EVALUATION

For each question, please CIRCLE THE NUMBER which best reflects your opinion. Based on the interview you have just given, how do you think the average hypothetical Stage 1 Psychology student would rate the person you have just interviewed on the following items:

1. Rate the contribution of the tutor to this course: Poor 1 2 3 4 5 6 7 Excellent

2. How would you characterize the tutor's ability to explain: Very poor 1 2 3 4 5 6 7 Excellent

3. Did the tutor seem well prepared for classes: No, never 1 2 3 4 5 6 7 Yes, always

4. The tutor's knowledge of course topics appeared to be: Inadequate 1 2 3 4 5 6 7 More than adequate

5. The tutor presented material at a level appropriate to class: Almost never 1 2 3 4 5 6 7 Almost always

6. The tutor was able to explain difficult material to my satisfaction: Almost never 1 2 3 4 5 6 7 Almost always

7. The tutor was able to answer questions clearly and concisely: Almost never 1 2 3 4 5 6 7 Almost always

8 The tutor stimulated my interest in the subject: Not at all 1 2 3 4 5 6 7 Very much

9. Did the tutor evaluate your work in a constructive and conscientious manner: Definitely no 1 2 3 4 5 6 7 Definitely

10. The tutor encouraged students to develop their own ideas and approaches to problems: Never 1 2 3 4 5 6 7 Frequently

11. The tutor initiated fruitful and relevant discussions: Never 1 2 3 4 5 6 7 Frequently

12. Was a good balance of student participation and tutor contribution achieved: No, very poor 1 2 3 4 5 6 7 Yes very good

13. Did the tutor try to involve all students in class activities: No, seldom 1 2 3 4 5 6 7 Yes, consistently

14. How often was discussion monopolized by only one or a few students: Almost never 1 2 3 4 5 6 7 Almost always

15. How often did the tutor encourage interaction among students: Never 1 2 3 4 5 6 7 Frequently

16. Was the tutor receptive to differing viewpoints or opinions: No, never 1 2 3 4 5 6 7 Yes always

17. How would you describe the tutor's attitude toward students in the course: Indifferent 1 2 3 4 5 6 7 Very helpful

18. How patient was the tutor in working with you: Very impatient 1 2 3 4 5 6 7 Very patient

19. How helpful was the tutor to students with problems: Not at all helpful 1 2 3 4 5 6 7 Very helpful

20. What the tutor expected of students was: Very unclear 1 2 3 4 5 6 7 Very clear

21. How accessible was the tutor to students outside class hours: Very inaccessible 1 2 3 4 5 6 7 Very accessible

22. The tutor was sensitive to student needs and concerns: Almost never 1 2 3 4 5 6 7 Almost always

Tutor Code: _____

Appendix 7

PSYCHOLOGY TUTOR EVALUATION

For each question, please CIRCLE THE NUMBER which best reflects your opinion. Based on your observation of the tutor's performance this year, how would you rate her on the following items:

1. Rate the contribution of the tutor to this course:	Poor 1 2 3 4 5 6 7 Excellent
2. How would you characterize the tutor's ability to explain:	Very poor 1 2 3 4 5 6 7 Excellent
3. Did the tutor seem well prepared for classes:	No, never 1 2 3 4 5 6 7 Yes, always
4. The tutor's knowledge of course topics appeared to be:	Inadequate 1 2 3 4 5 6 7 More than adequate
5. The tutor presented material at a level appropriate to class:	Almost never 1 2 3 4 5 6 7 Almost always
6. The tutor was able to explain difficult material to the student's satisfaction:	Almost never 1 2 3 4 5 6 7 Almost always
7. The tutor was able to answer questions clearly and concisely:	Almost never 1 2 3 4 5 6 7 Almost always
8 The tutor stimulated student's interest in the subject:	Not at all 1 2 3 4 5 6 7 Very much
9. Did the tutor evaluate student's work in a constructive and conscientious manner:	Definitely no 1 2 3 4 5 6 7 Definitely
10. The tutor encouraged students to develop their own ideas and approaches to problems:	Never 1 2 3 4 5 6 7 Frequently
11. The tutor initiated fruitful and relevant discussions:	Never 1 2 3 4 5 6 7 Frequently
12. Was a good balance of student participation and tutor contribution achieved:	No, very poor 1 2 3 4 5 6 7 Yes very good
13. Did the tutor try to involve all students in class activities:	No, seldom 1 2 3 4 5 6 7 Yes, consistently
14. How often was discussion monopolized by only one or a few students:	Almost never 1 2 3 4 5 6 7 Almost always
15. How often did the tutor encourage interaction among students:	Never 1 2 3 4 5 6 7 Frequently
16. Was the tutor receptive to differing viewpoints or opinions:	No, never 1 2 3 4 5 6 7 Yes always

17. How would you describe the tutor's attitude toward students in the course:	Indifferent 1 2 3 4 5 6 7 Very helpful
18. How patient was the tutor in working with students:	Very impatient 1 2 3 4 5 6 7 Very patient
19. How helpful was the tutor to students with problems:	Not at all helpful 1 2 3 4 5 6 7 Very helpful
20. What the tutor expected of students was:	Very unclear 1 2 3 4 5 6 7 Very clear
21. How accessible was the tutor to students outside class hours:	Very inaccessible 1 2 3 4 5 6 7 Very accessible
22. The tutor was sensitive to student needs and concerns:	Almost never 1 2 3 4 5 6 7 Almost always

Tutor Code: _____

Appendix 8

PSYCHOLOGY TUTOR EVALUATION

For each question, please CIRCLE THE NUMBER which best reflects your opinion.. Please consider each question separately and try not to let your overall reaction to the Tutor blind you to particular areas of strength or weakness.

1. Rate the contribution of the tutor to this course:	Poor 1 2 3 4 5 6 7 Excellent
2. How would you characterize the tutor's ability to explain?	Very poor 1 2 3 4 5 6 7 Excellent
3. Did the tutor seem well prepared for classes?	No, never 1 2 3 4 5 6 7 Yes, always
4. The tutor's knowledge of course topics appeared to be:	Inadequate 1 2 3 4 5 6 7 More than adequate
5. The tutor presented material at a level appropriate to class:	Almost never 1 2 3 4 5 6 7 Almost always
6. The tutor was able to explain difficult material to my satisfaction:	Almost never 1 2 3 4 5 6 7 Almost always
7. The tutor was able to answer questions clearly and concisely:	Almost never 1 2 3 4 5 6 7 Almost always
8 The tutor stimulated my interest in the subject:	Not at all 1 2 3 4 5 6 7 Very much
9. Did the tutor evaluate your work in a constructive and conscientious manner:	Definitely no 1 2 3 4 5 6 7 Definitely
10. The tutor encouraged students to develop their own ideas and approaches to problems:	Never 1 2 3 4 5 6 7 Frequently
11. The tutor initiated fruitful and relevant discussions:	Never 1 2 3 4 5 6 7 Frequently
12. Was a good balance of student participation and tutor contribution achieved?	No, very poor 1 2 3 4 5 6 7 Yes very good
13. Did the tutor try to involve all students in class activities?	No, seldom 1 2 3 4 5 6 7 Yes, consistently
14. How often was discussion monopolized by only one or a few students?	Almost never 1 2 3 4 5 6 7 Almost always
15. How often did the tutor encourage interaction among students?	Never 1 2 3 4 5 6 7 Frequently

16. Was the tutor receptive to differing viewpoints or opinions? — No, never 1 2 3 4 5 6 7 Yes always

17. How would you describe the tutor's attitude toward students in the course? — Indifferent 1 2 3 4 5 6 7 Very helpful

18. How patient was the tutor in working with you? — Very impatient 1 2 3 4 5 6 7 Very patient

19. How helpful was the tutor to students with problems? — Not at all helpful 1 2 3 4 5 6 7 Very helpful

20. What the tutor expected of students was: — Very unclear 1 2 3 4 5 6 7 Very clear

21. How accessible was the tutor to students outside class hours? — Very inaccessible 1 2 3 4 5 6 7 Very accessible

22. The tutor was sensitive to student needs and concerns: — Almost never 1 2 3 4 5 6 7 Almost always

Tutor Code: _____

Appendix 9

UNIVERSITY OF CANTERBURY DEPARTMENT OF PSYCHOLOGY

Person Perception Questionnaire

Please indicate your sex and age below:

SEX: male / female AGE: _____ years

Instructions

This questionnaire has been designed to investigate the different ways that people think about themselves and other people.

The questionnaire is ANONYMOUS so there is no need to put your name on it.

There are no RIGHT or WRONG answers. We are interested in your own PERCEPTIONS.

Please answer each question as HONESTLY and ACCURATELY as you can, but don't spend too much time thinking about each answer.

Scoring Procedure

The numbers on each scale represent the following degrees of agreement:

-3 = very untrue/inaccurate
-2 = moderately untrue/inaccurate
-1 = slightly untrue/inaccurate
Ø = neither true nor untrue, accurate nor inaccurate
1 = slightly true/accurate
2 = moderately true/accurate
3 = very true/accurate

Read each statement carefully and show your agreement or disagreement by circling ONE NUMBER on each scale. If what you believe or think about yourself is in agreement with the statement, circle one of the numbers from 1 to 3. If what you believe or think about yourself conflicts with the statement, circle one of the negative numbers. If you neither agree nor disagree circle the zero.

Example 1): If the statement said "I am very tall" you would circle the 2 if you were moderately tall, the zero if you were average height, and the -3 if you were very short.

Example 2): If the statement said "I believe in the death penalty for murder", you would circle the 3 if you very strongly believed that, the 1 if you slightly believed it, and the -3 if you strongly believed the opposite, i.e., murder should not carry the death penalty.

1) I have thought a lot about the family background and personal history of people who are close to me, in order to understand why they are the sort of people they are.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

2) I believe it is important to analyze and understand our own thinking processes.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

3) I think a lot about the influence that I have on other people's behavior.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

4) I have found that the relationships between a person's attitudes, beliefs, habits and character traits are usually simple and straightforward.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

5) If I see people behaving in a really strange or peculiar manner I usually put it down to the fact that they are strange people and don't bother to explain it any further.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

6) Once I have figured out a single cause for a person's behavior I don't usually go any further.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

7) I don't usually bother trying to analyze and explain people's behavior.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

8) I have often found that the basic cause for a person's behavior is located far back in time.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3	very true or accurate

9) I am very interested in understanding how my own thinking works when I make judgments about people or attach causes to their behavior.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

10) I think very little about the different ways that people influence each other.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

11) To understand a person's personality/behavior I have found it is important to know how that person's attitudes, beliefs, and character traits fit together.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

12) When I try to explain other people's behavior I concentrate on the person and don't worry too much about all the existing external factors that might be affecting them.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

13) I have found that the causes for people's behavior are usually complex rather than simple.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

14) I don't enjoy getting into discussions where the causes for people's behavior are being talked over.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

15) When I analyze a person's behavior I often find that the causes form a chain that go back in time, sometimes for years.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

16) I give little thought to how my own thinking works in the process of understanding or explaining people's behavior.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

17) I think very little about the influence that other people have on my behavior.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

18) I have thought a lot about the way that different parts of my personality influence other parts, e.g., beliefs affecting attitudes or attitudes affecting character traits.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

19) I think a lot about the influence that society has on other people.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

20) I usually find that complicated explanations for people's behavior are confusing rather than helpful.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

21) I really enjoy analyzing the reasons or causes for people's behavior.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

22) I have thought very little about my own family background and personal history in order to understand why I am the sort of person I am.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

23) When the reasons I give for my own behavior are different from someone else's, this often makes me think about the thinking processes that lead to my explanations.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

24) I believe that to understand a person you need to know and understand the people who that person has close contact with.

very untrue or inaccurate -3 -2 -1 Ø 1 2 3 very true or accurate

25) I tend to take people's behavior at face value and not worry about the inner causes for their behavior, e.g., attitudes, beliefs, etc.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3 very true or accurate

26) I think a lot about the influence that society has on my behavior and personality.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3 very true or accurate

27) I prefer simple rather than complex explanations for people's behavior.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3 very true or accurate

28) I am not really curious about human behavior.

very untrue or inaccurate	-3	-2	-1	Ø	1	2	3 very true or accurate